September, 1974 ESL-P-572

ANALYSIS OF BILINEAR STOCHASTIC SYSTEMS

by

Alan S. Willsky*, Steven I. Marcus**, and David N. Martin***

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

In this paper we consider the analysis of stochastic dynamical systems that involve multiplicative (bilinear) noise processes. After defining the systems of interest, we consider the evolution of the moments of such systems, the question of stochastic stability, and estimation for bilinear stochastic systems. Both exact and approximate methods of analysis are introduced, and, in particular, the uses of Lie-theoretic concepts and harmonic analysis are discussed.

(NASA-CR-141281) ANALYSIS OF BILINEAR STOCHASTIC SYSTEMS (Massachusetts Inst. of Tech.) 9 p HC $3.25 CSCL 12A

N75-15386

Unclas
G3/65 06537

* Assistant Professor of Electrical Engineering, M.I.T. The work of this author was supported by NSF under grant GK-42090 and by NASA under grant NGL-22-009-124.

** National Science Foundation Fellow.

*** Presently with Raytheon Company, Bedford, Mass.

ANALYSIS OF BILINEAR STOCHASTIC SYSTEMS

ALAN S. WILLSKY*, STEVEN I. MARCUS**, AND DAVID N. MARTIN***
Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

In this paper we consider the analysis of stochastic dynamical systems that involve multiplicative (bilinear) noise processes. After defining the systems of interest, we consider the evolution of the moments of such systems, the question of stochastic stability, and estimation for bilinear stochastic systems. Both exact and approximate methods of analysis are introduced, and, in particular, the uses of Lie-theoretic concepts and harmonic analysis are discussed.

I. INTRODUCTION

Recently, a great deal of effort has gone into developing a detailed theory for a class of nonlinear systems that possesses a great deal of structure -- the class of bilinear systems. The mathematical tools behind bilinear system analysis include not only many of the vector space techniques that are so valuable in linear system theory but also a number of tools drawn from the theories of Lie groups and differential geometry. It is the purpose of this paper to describe the several mathematical techniques that have been developed for the analysis of bilinear stochastic systems. The reader is ~~referred~~ to [1]-[7] for detailed descriptions of some of these techniques and for further references.

II. STOCHASTIC BILINEAR SYSTEMS

We consider the dynamical model

$$\dot{x}(t) = [A_0 + \sum_{i=1}^{N} u_i(t) A_i] x(t) \tag{1}$$

where the A_i are given n x n matrices, the u_i are scalar inputs, and x is either an n-vector or an n x n matrix. Let $\mathscr{L}$ be the Lie algebra generated by $A_0, \ldots, A_n$ -- i.e. the smallest subspace of n x n matrices containing $A_0, \ldots, A_n$ and closed under the commutator product $[P,Q] = PQ-QP$--and let G be the Lie group associated with $\mathscr{L}$ - - i.e. the smallest group (under matrix multiplication) containing $\exp(L)\ \forall\, L \in \mathscr{L}$. It is well known that for piecewise continuous u_i, we have $x(t) \in G\, x(0)$, t. If the u_i are white noise processes, we have that $x \in Gx(0)$ holds almost surely for the solution to the Ito equation

$$dx(t) = \{[A_0 + \frac{1}{2} \sum_{i,j=1}^{N} R_{ij}(t) A_i A_j] dt + \sum_{i=1}^{N} A_i dv_i(t)\} x(t) \tag{2}$$

where v is a Brownian motion with $E[dv(t)dv'(t)] = R(t)dt$. If u is smoother than white noise - - for instance if u satisfies

---

* Assistant Professor of Electrical Engineering, M.I.T. The work of this author was supported by NSF under grant GK-42090 and by NASA under grant NGL-22-009-124.

** National Science Foundation Fellow.

*** Presently with Raytheon Company, Bedford, Mass.

$$d\xi(t) = F(t)\xi(t)dt + G(t)dw(t) + \alpha(t)dt, \quad u(t) = H(t)\xi(t) \tag{3}$$

where w is a Brownian motion -- we needn't include the correction term contained in (2). In this case x is not a Markov process, but the pair (x,ξ) is; however, when we regard (x,ξ) as the state, the system (1),(3) involves products of state variables. As we shall see, the analysis in this case is more difficult than in the white noise case.

III. MOMENT EVOLUTION EQUATIONS

3.1 Systems Driven by White Noise

In this subsection we present moment equations derived by Brockett [4]. Consider (2) where, for simplicity, we assume that $R(t) = I$. Recall that the number of linearly independent homogeneous polynomials of degree p in n variables (i.e. $f(cx_1,\dots,cx_n) = c^p f(x_1,\dots,x_n)$) is

$$N(n,p) = \binom{n+p-1}{p} \tag{4}$$

We choose a basis for this space of polynomials consisting of the elements

$$\sqrt{\binom{p}{p_1}\binom{p-p_1}{p_2}\cdots\binom{p-p_1-\dots-p_{n-1}}{p_n}}\; x_1^{p_1}\cdots x_n^{p_n}\;;\; \sum_{i=1}^{n} p_i = p,\; p_i \geq 0 \tag{5}$$

If we denote the vector of these elements (ordered lexicographically) by $x^{[p]}$, then $||x||^p = ||x^{[p]}||$, where $||x||^2 = x'x$. It is clear that if x satisfies $\dot{x} = Ax$, then $x^{[p]}$ satisfies $\dot{x}^{[p]} = A^{[p]}x^{[p]}$. We regard this as the definition of $A^{[p]}$, which is closely related to Kronecker sum matrices [10],[11].

The Stratonovich analog of (2) is

$$dx(t) = [A_0 dt + \sum_{i=1}^{N} A_i dv_i(t)]x(t) \tag{6}$$

Since Stratonovich equations obey the rules of ordinary calculus, we obtain

$$dx^{[p]}(t) = \left[A_0{}^{[p]}dt + \sum_{i=1}^{N} A_i{}^{[p]}dv_i(t)\right]x^{[p]}(t) \tag{7}$$

Using the moment equations for Stratonovich equations, we have an evolution for the pth moments of (2):

$$\frac{d}{dt}E(x^{[p]}(t)) = \{A_0{}^{[p]} + \frac{1}{2}\sum_{i=1}^{N}(A_i{}^{[p]})^2\}E(x^{[p]}(t)) \tag{8}$$

3.2 Systems Driven by Colored Noise: Exact Expressions

Definition 1: We associate with any Lie algebra $\mathscr{L}$ its derived series

$$\mathscr{L}^{(0)} = \mathscr{L}$$

$$\mathscr{L}^{(n+1)} = [\mathscr{L}^{(n)},\mathscr{L}^{(n)}] = \{[A,B] \mid A,B \in \mathscr{L}^{(n)}\},\; n \geq 0 \tag{9}$$

$\mathscr{L}$ is abelian if $\mathscr{L}^{(1)} = \{0\}$ and is solvable if $\mathscr{L}^{(n)} = \{0\}$ for some n.

Theorem 1: A matrix Lie algebra $\mathscr{L}$ is solvable if and only if there exists a nonsingular matrix P (possibly complex-valued) such that PAP^{-1} is upper triangular (zero below diagonal) for all $A \in \mathscr{L}$.

Proof: See [13].

We now consider (1) where we assume that u is a zero mean Gaussian random process with $E[u(t)u'(s)] = P(t,s)$. Let $\mathscr{L}$ be the Lie algebra generated by $A_0, A_1, \ldots, A_N$. We claim that one can construct closed form expressions for $E[x^{[p]}(t)]$ if $\mathscr{L}$ is solvable. Assuming that L is solvable (and that x(0) is independent of u), we can obtain a closed form expression for Φ_u, the state transition matrix for (1), as a function of u. First we find the matrix P as in Theorem 1 such that each $B_i = PA_iP^{-1}$ is upper triangular. Then we can solve the equation

$$\dot{\Psi}_u(t,0) = [B_0 + \sum_{i=1}^{N} B_i u_i(t)]\Psi_u(t,0), \quad \Psi_u(0,0) = I \tag{10}$$

by straightforward calculations (see the example in Subsection 4.2), and we compute

$$x(t) = P^{-1}\Psi_u(t,0)P\, x(0) \tag{11}$$

and Ψ_u involves nothing more complicated than exponentials of integrals of components of u, polynomials in u, and various combinations, products, and integrals of such ~~quan~~tities. Since u is Gaussian, we can evaluate the expectations of such quantities in closed form (see Subsection 4.2) and can thus obtain a closed form expression for $E[x(t)]$. We can also obtain closed form expressions for $E[x^{[p]}(t)]$, as $x^{[p]}$ consists of the same types of functionals of u (see [7]).

3.3 Systems Driven by Colored Noise--Approximate Analysis

Consider the system (1) where u is given by (3). In this section we describe an approximate method that can be used for moment analysis if L is not solvable. This approach is based on the truncation of the cummulants of a random process [14] and can be used for arbitrary nonlinear systems, although it is particularly well-suited to systems with polynomial nonlinearities. We will illustrate the approach by example and refer the reader to [1], [2], and [9] for more detailed discussions.

Consider the scalar example

$$dx(t) = \alpha x^2(t)dt + \beta x(t)dw(t) \tag{12}$$

where $E[dw^2(t)] = dt$. We compute [12]

$$\dot{m}_1(t) = \alpha m_2(t), \quad \dot{m}_2 = 2\alpha m_3(t) + \beta^2 m_2(t)$$
$$\dot{m}_3(t) = 3\alpha m_4(t) + 3\beta^2 m_3(t) \tag{13}$$

where $m_i = E(x^i)$.

From these equations it is clear that any attempt to study the properties of (13) must involve a truncation of the infinite set of coupled equations.

As discussed in [1] and [14], the direct truncation method -- setting to zero all moments greater than some given order -- can cause difficulties, but the truncation of the cumulants -- the coefficients of the Taylor series expansion of the logarithm of the characteristic function -- is quite reasonable (for instance, only the first two cumulants are nonzero for

Gaussian densities). By comparing terms in the series expansions for the characteristic function, one can compute relationships between the cumulants and moments. For instance, the fourth cumulant satisfies

$$k_4 = m_4 - 3m_2^2 - 4m_1m_3 + 12m_1^2m_2 - 6m \qquad (14)$$

Thus if we assume $k_4 = 0$ and use (14) to express m_4 in terms of m_i, $i \leq 3$, we can use this approximation in the moment equations (13).

Finally, we note that the cumulants technique can be used to design suboptimal estimators for bilinear systems (see [2])

IV. STOCHASTIC STABILITY

A problem of considerable theoretical and practical interest is the question of the stability of linear systems containing multiplicative noise processes. As discussed in [7] and [8], such systems often arise in linear feedback systems in which the actuator, sensor, and feedback gains may be stochastic in nature. In this case, a model such as (1) arises, in which the u_i represent the random elements in the system.

Definition 2: A vector random process x is pth order stable if $E[x^{[p]}(t)]$ is bounded for all t, and x is pth order asymptotically stable if

$$\lim_{t\to\infty} E[x^{[p]}(t)] = 0 \qquad (15)$$

The systems (1),(3) and (2) are pth order (asymptotically) stable if the solution x is pth order (asymptotically) stable for all initial conditions $x(0)$ that are independent of ξ (in the (1),(3) case) or the v_i (in (2)) and such that $E[x^{[p]}(0)] < \infty$.

4.1 The White Noise Case

Consider the system (2). Using (8), we have the following.

Theorem 2: The system (2) with R=I is pth order asymptotically stable if and only if the matrix

$$D_p = A_0^{[p]} + \frac{1}{2}\sum_{i=1}^{N} (A_i^{[p]})^2 \qquad (16)$$

has all its eigenvalues in the left-half plane ($\mathrm{Re}(\lambda) < 0$). The system is pth order stable if D_p has all its eigenvalues in $\mathrm{Re}(\lambda) \leq 0$ and if the eigenvalues with real parts $= 0$ have all higher order residues $= 0$ (i.e. there are no $te^{\lambda t}$ terms in $e^{D_p t}$ with $\mathrm{Re}(\lambda) = 0$; see [15, p.55]).

For examples of this result, we refer the reader to [7].

4.2. The Colored Noise Case

We have seen in Subsection 3.2 that we can obtain closed form moment expressions for systems of the form given in (1) when u is a Gaussian colored noise process and $\mathscr{L}$ is solvable. We can use these expressions to obtain stochastic stability criteria. We illustrate these ideas by means of an example (see [7] for other examples).

Example 1: Consider the system

$$\dot{x}(t) = [A_0 + A_1\xi_1(t) + A_2\xi_2(t)]x(t) \qquad (17)$$

$$A_0 = \alpha \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \tag{18}$$

with ξ_1, ξ_2 zero mean and

$$E\left[\xi_i(t)\,\xi_i(t+\tau)\right] = \frac{1}{2}\left[\sigma_1^2 e^{-k_1|\tau|} + \sigma_2^2 e^{-k_2|\tau|}\right], \quad i=1,2 \tag{19}$$

$$E\left[\xi_1(t)\,\xi_2(t+\tau)\right] = \frac{1}{2}\left[\sigma_1^2 e^{-k_1|\tau|} - \sigma_2^2 e^{-k_2|\tau|}\right]. \tag{20}$$

In this case L is solvable, and defining $y = Px$, with

$$P = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \tag{21}$$

we have

$$y(t) = \begin{bmatrix} \eta_1(t,0) & \int_0^t \eta_1(t,s)\eta_2(s,0)\,[\alpha + \xi_1(s)]\,ds \\ 0 & \eta_2(t,0) \end{bmatrix} \tag{22}$$

$$\eta_i(t_1,t_2) = \exp \; \alpha(t_2-t_1) + \int_{t_1}^{t_2} \xi_i(s)\,ds \tag{23}$$

Some lengthy but straightforward calculations yield the following necessary and sufficient condition for first order asymptotic stability

$$\alpha < -\frac{1}{2}\left(\frac{\sigma_1^2}{k_1} + \frac{\sigma_2^2}{k_2}\right) \tag{24}$$

and the system is first order stable if equality holds in (24).

V. ESTIMATION FOR BILINEAR SYSTEMS

There have been a number of results obtained on estimation for bilinear systems. In this section we illustrate two classes of results. The first of these involves estimation on nilpotent Lie groups. For such systems one can obtain finite dimensional optimal nonlinear filters [3], and we illustrate the result for the 3-dimensional case in Subsection 5.1. An approximate technique based on the use of harmonic analysis on compact Lie groups is illustrated in Subsection 5.2. For more detailed discussions of bilinear estimation, we refer the reader to [2],[3],[5], and [6].

5.1 Estimation on Nilpotent Lie Groups

Definition 3: A Lie algebra L is called *nilpotent* if the sequence of Lie algebras

$$\mathscr{L}^0 = \mathscr{L}$$

$$L^{n+1} = [L, L^n] = \{[A,B] \mid A \in L,\ B \in L^n\},\ n \geq 0 \tag{25}$$

terminates in zero. Note that abelian => nilpotent => solvable, but none of the reverse implications hold in general.

Consider the system (1) with $A_0 = 0$, $N=3$, and

$$A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{26}$$

It is easy to see that these matrices generate a nilpotent Lie algebra. Also, suppose that u satisfies (3) with (for simplicity) $H = I$ and $\alpha = 0$, and let x be the 3x3 transition matrix for (1) (i.e. $x(0) = I$). Then

$$x(t) = \begin{bmatrix} 1 & \int_0^t \xi_1(\sigma)d\sigma & \int_0^t \xi_2(\sigma)d\sigma + \int_0^t \int_0^{\sigma_1} \xi_1(\sigma)\xi_3(\sigma_2)d\sigma_2 d\sigma_1 \\ 0 & 1 & \int_0^t \xi_3(\sigma)d\sigma \\ 0 & 0 & 1 \end{bmatrix} \tag{27}$$

Suppose we have the linear observations

$$dz(t) = C(t)\xi(t)dt + R^{1/2}(t)dv(t) \tag{28}$$

where v is a Brownian motion process and $R > 0$. Then the conditional mean, $\hat{x}(t|t) = E(x(t)|z(s), 0 \leq s \leq t)$, can be generated by a finite dimensional nonlinear filter as follows (see [2],[3] for details): let

$$P(\sigma_1,\sigma_2,t) = E[(\xi(\sigma_1) - \hat{\xi}(\sigma_1|t))(\xi(\sigma_2) - \hat{\xi}(\sigma_2|t))'|z(s), 0\leq s\leq t] \tag{29}$$

Then $P(\sigma,t,t) = k(t,\sigma)\Sigma(t)$, and is nonrandom, where

$$\dot{\Sigma} = F\Sigma + \Sigma F' + GG' - \Sigma C'R^{-1}C\Sigma \tag{30}$$

$$K'(t,\sigma) = -[F'(t) + \Sigma^{-1}(t)G(t)G'(t)]K'(t,\sigma) \quad ; \quad K(\sigma,\sigma) = I \tag{31}$$

Consider the additional state variables

$$\dot{\eta}(t) = \xi(t) \tag{32}$$

$$\dot{\alpha}(t) = \left[\int_0^t K_3'(t,\sigma)e_1'd\sigma \right] \xi(t) - [F'(t) + \Sigma^{-1}(t)G(t)G'(t)]\alpha(t) \tag{33}$$

$$\dot{\beta}(t) = [e_1e_3']\eta(t) - [F'(t) + \Sigma^{-1}(t)G(t)G'(t)]\beta(t) \tag{34}$$

where $\eta(0) = \alpha(0) = \beta(0) = 0$, K_j is the jth row of K, and e_i is the ith unit vector in R^3. We implement a Kalman filter for (3), (28), (32)-(34), and then

$$\hat{x}(t|t) = \begin{bmatrix} 1 & \hat{\eta}_1(t|t) & \hat{\eta}_2(t|t) + \hat{\gamma}(t|t) \\ 0 & 1 & \hat{\eta}_3(t|t) \\ 0 & 0 & 1 \end{bmatrix} \tag{35}$$

where $\hat{\gamma}(0|0) = 0$ and

$$\begin{aligned} d\hat{\gamma}(t|t) = {} & \left[\hat{\xi}_1(t|t)\hat{\eta}_3(t|t) + \int_0^t \Sigma_{13}(t,\sigma,t)\,d\sigma\right]dt \\ & + [\hat{\alpha}'(t|t)+\hat{\beta}(t|t)]\,\Sigma(t)C'(t)R^{-1}(t)\,[dz(t)-C(t)\hat{\xi}(t|t)dt] \end{aligned} \tag{36}$$

The key to this result is the fact that K is a state transition matrix with a finite-dimensional representation (31). Also, as mentioned earlier, this result can be extended to arbitrary nilpotent Lie groups, some solvable Lie groups, and to other nonlinear systems. The key to these extensions is the fact that products and transposes of weighting patterns with finite-dimensional realizations also have finite-dimensional realizations (see [3]).

5.2 The Use of Harmonic Analysis in Suboptimal Filter Design

In this subsection we illustrate the use of harmonic analysis in bilinear estimation by studying a phase tracking problem with the aid of Fourier series analysis. For extensions of these results to other problems, including attitude estimation with the aid of spherical harmonic analysis, we refer the reader to [2]. Suppose we receive the signal z defined by

$$dz(t) = [\sin\theta(t)]dt + r^{1/2}(t)dw(t) \tag{37}$$

$$d\theta(t) = w_c dt + q^{1/2}(t)dv(t) \tag{38}$$

where v and w are independent standard Brownian motions independent of the random initial phase $\theta(0)$. We desire to track the signal phase, θ mod 2π. As discussed in [6], a useful estimation criterion is to choose $\tilde{\theta}(t)$ to minimize $E[1-\cos(\theta(t)-\tilde{\theta}(t))|z(s), 0 \le s \le t]$.

Putting the problem into Cartesian coordinates by defining $x_1 = \sin\theta$, $x_2 = \cos\theta$, we have the bilinear equations

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} -q(t)dt/2 & (w_c dt+q^{1/2}(t)dv(t)) \\ -(w_c dt+q^{1/2}(t)dv(t)) & -q(t)dt/2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{39}$$

$$dz(t) = x_1(t)dt + r^{1/2}(t)dw(t) \tag{40}$$

and the optimal estimate is

$$\tilde{\theta}(t) = \tan^{-1} \frac{\hat{x}_1(t|t)}{\hat{x}_2(t|t)} \tag{41}$$

We note that $\tilde{x}_1 = \sin\tilde{\theta}$, $\tilde{x}_2 = \cos\tilde{\theta}$ are minimum variance estimates of x_1 and

x_2 under the constraint $\tilde{x}_1^2 + \tilde{x}_2^2 = 1$.

To solve this problem, we expand the conditional probability density for $\theta(t)$ given $\{z(s),\ 0 \leq s \leq t\}$ in a Fourier series. The coefficients are

$$c_n(t) = \frac{1}{2\pi} E[e^{-in\theta(t)} | z(s),\ 0 \leq s \leq t] = b_n(t) - ia_n(t) \tag{42}$$

We then have the coupled equations [6]

$$dc_n = -[inw_c + \frac{n^2}{2} q] c_n dt + \frac{(c_{n-1} - c_{n+1})}{2i} + 2\pi c_n I_m(c_1) \left[\frac{dz + 2\pi I_m(c_1) dt}{r}\right] \tag{43}$$

See [6] for a discussion of the structure of the optimal filter. Note that $x_2 - ix_1 = 2\pi c_1$, and thus we need only compute c_1 to determine $\hat{\theta}$; however, c_2 couples into c_1, c_3 into c_2, etc. Thus we need to approximate the optimal filter to obtain an implementable system. One approach considered in [6] is to make an approximation based on the assumption that the conditional density is a "folded normal density," the circular analog of the normal density. The coefficients of such a density depend on only two parameters

$$2\pi c_n = e^{-n^2\gamma/2} e^{-in\eta} \tag{44}$$

and $c_2 = 8\pi^3 |c_1|^2 c_1^2$. Using this relationship as an approximation in (44), we obtain an equation for c_1 that is uncoupled from the other coefficients. We refer the reader to [6] for simulation results that indicate that this filter performs better than the optimal phase-lock loop. In addition, directly analogous results using spherical harmonics have been obtained in [2] for the estimation of processes on the sphere.

VI. CONCLUSIONS

In this paper we have described a number of techniques for the analysis of bilinear stochastic systems. As can be seen from our results and from those in the references, this is a class of systems that is rich in both structure and practical applications.

References

1. A.S. Willsky and S.I. Marcus, "Analysis of Bilinear Noise Models in Circuits and Devices," Monograph of the Colloquium on the Application of Lie Group Theory to Nonlinear Network Problems, 1974 IEEE International Symposium on Circuits and Systems, San Francisco, Calif., April 1974.

2. A.S. Willsky and S.I. Marcus, "Estimation for Bilinear Stochastic Systems," presented at the U.S.-Italy Seminar on Variable Structure Systems, Portland, Oregon, May 26-31, 1974; also M.I.T. Electronic Systems Laboratory Report ESL-R-544.